AD-A059 657 CALIFORNIA UNIV DAVIS DEPT OF MATHEMATICS F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION FOR A CLASS OF MULTINOMIAL DISTRI--ETC(U)
JUL 78 F J SAMANIEGO, L E JONES AFOSR-77-3180
UNCLASSIFIED TR-6 AFOSR-TR-78-1267 NL

| OF |
AD
A059657

END
DATE
FILMED
12-78
DDC

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0 4.5 5.0 5.6 6.3

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

TECHNICAL REPORT SERIES
ON
SIGNAL PLUS NOISE MODELS

Department of Mathematics
University of California
Davis, California

Maximum Likelihood Estimation for a Class of Multinomial Distributions Arising in Reliability

F. J. Samaniego and L. E. Jones

Technical Report No. 6

July 1978

Department of Mathematics
University of California
Davis, California

This research was supported in part by the Air Force Office of Scientific Research under contract AFOSR-77-3180.

SUMMARY

Let X_i, $i=1,\ldots,k$ be independent Bernoulli variables, with $X_i \sim B(1,p_i)$. Let $Y = \Sigma X_i$, and consider a multinomial experiment based on n independent, identically distributed observations $Y_1,\ldots,Y_n$. This model is identifiable in the ordered parameter vector $(p_{(1)},\ldots,p_{(k)})$, where $p_{(i-1)} \leq p_{(i)}$, and arises in reliability experiments in which k components in parallel have potentially different probabilities of failure. The family of multinomial distributions with the structure described above is properly contained in the family of general multinomial distributions with $(k+1)$ classes. Maximum likelihood estimation for this family is considered, and it is shown that for sufficiently large n, the maximum likelihood estimate of $(p_{(1)},\ldots,p_{(k)})$ may be identified with high probability from the roots of a k^{th} degree polynomial whose coefficients are consistent estimates of the elementary symmetric functions of the ratios $\theta_{(i)} = p_{(i)}/(1-p_{(i)})$. A simulation study for the case $k = 2$ sheds light on the sample size required.

I. INTRODUCTION

Let X_i, $i=1,\ldots,k$ be independent Bernoulli random variables with potentially different probabilities of success p_i, $i=1,\ldots,k$. We denote this situation by $X_i \sim B(1,p_i)$, $i=1,\ldots,k$. Let $Y = \Sigma_{i=1}^{k} X_i$, and assume that a random sample $Y_1,Y_2,\ldots,Y_n$ is available. The common distribution of these Y's is the k-fold convolution to be denoted $\overset{k}{\underset{i=1}{*}} B(1,p_i)$.

sub i — is denoted — approx — sub i — sum over i = 1 to k of X sub i — *(i = 1 to k) — sub i

This note concerns the estimation of the parameters of this convolution based on the Y sample via the method of maximum likelihood.

Estimation problems for convolution models have been considered by several authors. For example, Gaffey (1959) constructed a consistent estimator for the distribution of one component of a continuous convolution model under the assumption that the distribution of the second component was known. Sclove and Van Ryzin (1969) derived method of moments estimators for a variety of convolution models. Maximum likelihood estimation has met substantial resistance for convolution models due to the cumbersome nature of the likelihood function, which, for discrete components, consists of a product of sums of products of component probabilities. Samaniego (1976) used a characterization of convoluted Poisson distributions to facilitate maximum likelihood estimation of the Poisson parameter. A similar approach was taken by Samaniego (1977) for maximum likelihood estimation in convoluted binomial distributions. Both of these studies have dealt with one-parameter models in which one component of the convolution has a known distribution. In general, maximum likelihood estimation for multiparameter problems (with the exception of the problem treated here) has as yet proven untractable.

The convolution $\overset{k}{\underset{1}{*}} B(1,p_i)$ is not well defined for estimation purposes, since the model is not identifiable in the parameter vector $\underset{\sim}{p} = (p_1,\ldots,p_k)$. It is clear that any permutation of the components of $\underset{\sim}{p}$ gives rise to the same distribution. It is easy to verify that this is precisely the extent of multiplicity in the model, and that the model is therefore identifiable in the ordered parameter vector $(p_{(1)},p_{(2)},\ldots,p_{(k)})$, where $p_{(i-1)} \leq p_{(i)}$ $\forall i$. We will consider estimation of the ordered parameter vector.

The model $\overset{k}{\underset{1}{*}} B(1,p_i)$ arises naturally in reliability experiments. Suppose k components are operating independently in an r out of k system, and their probabilities of operating successfully over a specified

time period are p_i, $i=1,\ldots,k$. The four tires of an automobile yield an example of components whose life lengths are independent, but are not identically distributed because of the different stresses experienced due to tire location. An r out of k system is only as reliable as its r^{th} best component, and thus it may be of interest to estimate the ordered parameter vector. The model we deal with assumes that no subsystem information is available, that is, the X variables from which the observable Y is constructed are themselves unobservable. Such an assumption is realized in many biological or engineering systems.

A number of authors of reliability texts have discussed the model $\overset{k}{\underset{1}{*}} B(1,p_i)$ as an introductory example (see, for instance, Barlow and Proschan (1975) p. 20 ff.). While many properties of the model are quite well known, inference questions remain largely uninvestigated. The estimation problem at hand is tangentially related to the estimation of parameters of mixtures of binomial distributions studied by Blischke (1964), and to estimation under order restrictions studied extensively in Barlow et al. (1972). The problem does not seem to benefit, however, from either of the approaches used in these studies. An estimation problem for this model with $k = 2$ was considered by Buehler (1957). In that paper, subsystem information was assumed available and a minimum width confidence interval for the reliability p_1p_2 of a series system was obtained. Our own interest in the problem considered here originated from an attempt to derive the maximum likelihood estimate of the parameter vector (p_1,p_2) in the convolution $B(N,p_1) * B(M,p_2)$ of two binomial distributions. It is interesting that this model is subsumed by the model $\overset{k}{\underset{1}{*}} B(1,p_i)$ for $k = N + M$, so that the comments developed here in fact apply to the binomial convolution. The approach

taken here is inefficient for the latter model, however, because the special structure of the binomial convolution is ignored by this approach.

II. THE CASE k = 2

We summarize in this section the derivation of the maximum likelihood estimates of the ordered parameters $p_{(1)}$ and $p_{(2)}$ in a two-component system. The character of the general problem can be gleaned from this case. Moreover, the solution for k = 2 is complete, whereas in the general problem, we consider only certain important special cases. For the case with k = 2, let n_i denote the observed frequency of the event Y = i, for i=0,1,2, in the sample $Y_1,\ldots,Y_n$. The likelihood function is given by

$$L(n_0,n_1,n_2,p_1,p_2)$$

$$= \frac{n!}{n_0!n_1!n_2!}\Big((1-p_1)(1-p_2)\Big)^{n_0}\Big(p_1(1-p_2) + p_2(1-p_1)\Big)^{n_1}\Big(p_1p_2\Big)^{n_2}.$$

The maximum likelihood estimate of $(p_{(1)},p_{(2)})$ is obtained from separate maximization problems for various possible data configurations, and is displayed in the table below:

TABLE I: MLE for $k = 2$

	Case	MLE $\left(\hat{p}_{(1)}, \hat{p}_{(2)}\right)$
1.	$n_0 = n$	$(0,0)$
2.	$n_1 = n$	$(0,1)$
3.	$n_2 = n$	$(1,1)$
4.	$n_0 = 0,\ n_1 \cdot n_2 > 0$	$\left(\frac{n_2}{n_1+n_2}, 1\right)$
5.	$n_1 = 0,\ n_0 \cdot n_2 > 0$	$\left(\frac{n_2}{n_0+n_2}, \frac{n_2}{n_0+n_2}\right)$
6.	$n_2 = 0,\ n_0 \cdot n_1 > 0$	$\left(0, \frac{n_1}{n_0+n_1}\right)$
7.	$\Pi\, n_i > 0,\ n_1^2 < 4n_0 \cdot n_2$	$\left(\frac{n_1+2n_2}{2n}, \frac{n_1+2n_2}{2n}\right)$
8.	$\Pi\, n_i > 0,\ n_1^2 \geq 4n_0 \cdot n_2$	$\frac{n_1+2n_2}{2n} \pm \frac{1}{2n}\sqrt{n_1^2 - 4n_0 \cdot n_2}$

In this estimation problem, the MLE is restricted to the simplex in the plane bounded by the lines $p_{(1)} = 0$, $p_{(2)} = 1$ and $p_{(1)} = p_{(2)}$. The first seven cases in Table I are boundary solutions. Even in the straightforward problem examined in this section, however, boundary solutions require some work. For example, in Case 4 above, the likelihood maximized on the boundary $p_{(2)} = 1$ takes on the value

$$L_1 = \left(\frac{n_1}{n_1+n_2}\right)^{n_1}\left(\frac{n_2}{n_1+n_2}\right)^{n_2},$$

while the likelihood maximized on the boundary $p_{(1)} = p_{(2)}$ takes on the value

$$L_2 = \left(\frac{n_1}{n_1+n_2}\right)^{n_1}\left(\frac{n_1+2n_2}{2n_1+2n_2}\right)^{n_1+2n_2}.$$

The fact that $L_1 > L_2$ may be inferred from the fact that the function $(1 + y/x)^x$, for fixed y, increased to e^y as x increases from zero to infinity. In the general problem, the simplex over which the likelihood is maximized is bounded by a multitude of hyperplanes, and boundary searches for the MLE are at the very least quite tedious.

Let us view the estimation problem from another perspective. Suppose we look for the MLE in terms of the basic multinomial probabilities (P(Y=0), P(Y=1)) subject to the constraints imposed by the model. We illustrate the outcome in the figures below, in which the dot represents the location of the unconstrained MLE $(n_0/n, n_1/n)$ and the box represents (albeit oversimplified) the constrained parameter space.

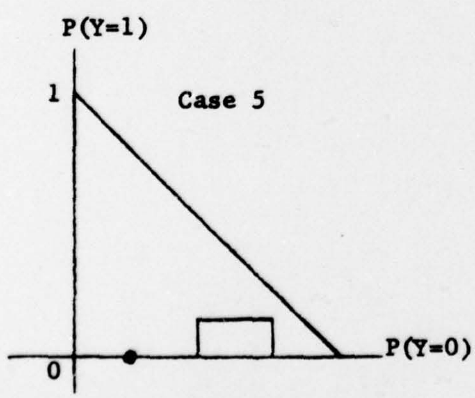

FIGURE I

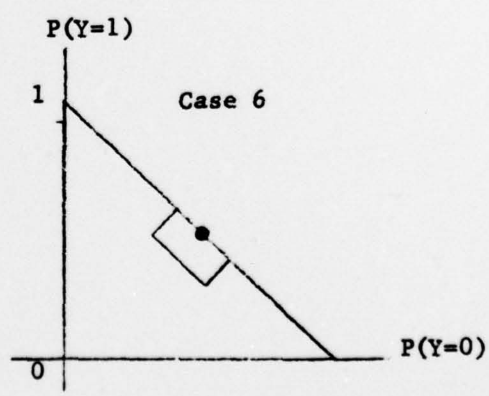

FIGURE II

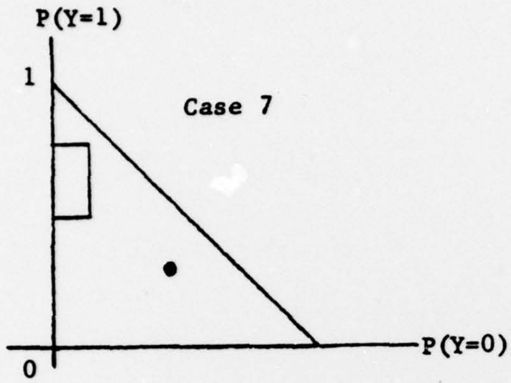

FIGURE III

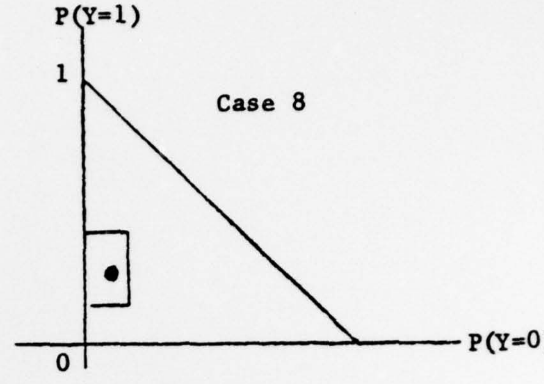

FIGURE IV

We make several observations. First, we notice that the unconstrained MLE may lie on the boundary or in the interior of the parameter space, and may be inside or outside of the constrained space. When it lies outside the constrained space, it is easy to argue that the constrained MLE will lie on the boundary of the constrained space. When the unconstrained MLE lies within the constrained space, the MLE for $(p_{(1)},p_{(2)})$ may be obtained, at least in theory, by the invariance property of MLE's. Our final observation is one which we will make more precise in our treatment of the general problem. We note that among the seven cases in which the nonzero n_i are consecutive integers, only in Case 7 does the unconstrained MLE lie outside of the constrained space. We thus conclude that the MLE for $(p_{(1)},p_{(2)})$ may be found by the invariance property of MLE's in almost all cases in which the integers i for which n_i is nonzero are consecutive.

III. THE GENERAL CASE

Let $Y_1,\ldots,Y_n$ be i.i.d. according to the distribution $\overset{k}{\underset{i=1}{*}} B(1,p_i)$, and let n_i be the observed frequency of the event $Y = i$ for $i=0,1,\ldots,k$.

The likelihood function may be written as

$$L(\underset{\sim}{n},\underset{\sim}{p}) = \frac{n!}{\prod_{i=0}^{k} n_i!}\left(\prod_{i=1}^{k}(1-p_i)\right)^{n_0}\left(\sum_{i=1}^{k} p_i \prod_{j\neq i}(1-p_j)\right)^{n_1}\cdots\left(\prod_{i=1}^{k} p_i\right)^{n_k} \tag{3.1}$$

Maximizing L with respect to $\underset{\sim}{p}$ is a difficult problem for several reasons. First, there are $2^{k+1} - 1$ different data configurations (where certain n_i are zero and the rest are positive), and the function L to be maximized is different for each. Secondly, the likelihood equations $\frac{\partial}{\partial p_i} L = 0$, $i=1,\ldots,k$, form a system of k equations each of which involves all k parameters in a nontrivial way. Another approach to maximum likelihood estimation is maximization of L with respect to the multinomial probabilities $\{P(Y=i)\}$, subject to the constraints on these probabilities imposed by the model. The difficulty with Lagrangian maximization in this problem is that the constraints are extremely complex and, for practical purposes, defy description. We will pursue a third approach, one that cannot be guaranteed to produce the MLE for any fixed sample size n, but which produces the MLE with limiting probability one when the parameters $p_{(i)}$, $i=1,\ldots,k$ are distinct.

We first consider data configurations for which $\prod_{i=0}^{k} n_i > 0$. We attempt to find the MLE for the ordered parameter vector $(p_{(1)},\ldots,p_{(k)})$ by the invariance property of MLE's, that is, by solving the system of equations

$$\begin{aligned}
\frac{n_0}{n} &= \hat{P}(Y=0) = \prod_{i=1}^{k}(1-\hat{p}_{(i)}) \\
\frac{n_1}{n} &= \hat{P}(Y=1) = \sum_{i=1}^{k}\hat{p}_{(i)}\prod_{j\neq i}(1-\hat{p}_{(j)}) \\
&\vdots \\
\frac{n_k}{n} &= \hat{P}(Y=k) = \prod_{i=1}^{k}\hat{p}_{(i)}
\end{aligned} \tag{3.2}$$

The system (3.2) consists of k+1 equations, but is determined by any k of them. We divide the last k equations by the first to obtain an equivalent system

$$
\begin{aligned}
\sum_{i=1}^{k} \left(\frac{\hat{p}_{(i)}}{1-\hat{p}_{(i)}}\right) &= \frac{n_1}{n_0} \\
\sum_{i<j} \left(\frac{\hat{p}_{(i)}}{1-\hat{p}_{(i)}}\right)\left(\frac{\hat{p}_{(j)}}{1-\hat{p}_{(j)}}\right) &= \frac{n_2}{n_0} \\
&\vdots \\
\sum_{i=1}^{k} \left(\prod_{j\neq i} \frac{\hat{p}_{(j)}}{1-\hat{p}_{(j)}}\right) &= \frac{n_{k-1}}{n_0} \\
\prod_{i=1}^{k} \left(\frac{\hat{p}_{(i)}}{1-\hat{p}_{(i)}}\right) &= \frac{n_k}{n_0}
\end{aligned}
\tag{3.3}
$$

We recognize the left hand side of (3.3) as the elementary symmetric functions in $\left\{\frac{\hat{p}_{(i)}}{1-\hat{p}_{(i)}}\right\}$, which implies that if the system (3.3) has a solution $(p_{(1)},\ldots,p_{(k)})$, it is unique and may be obtained as

$$\hat{p}_{(i)} = \frac{\hat{\theta}_{(i)}}{1+\hat{\theta}_{(i)}} \qquad i=1,\ldots,k \tag{3.4}$$

where $\hat{\theta}_{(i)}$, $i=1,\ldots,k$ are the ordered roots of the polynomial

$$p(x) = \sum_{i=0}^{k} (-1)^i n_i x^{k-i} \ . \tag{3.5}$$

The maximum likelihood estimate of $(p_{(1)},\ldots,p_{(k)})$ may be identified from (3.4) only when the polynomial $p(x)$ has k nonnegative roots. It is of

course possible for $p(x)$ to have some complex roots, or for some of the roots of $p(x)$ to be negative. However, one can show that when $0 < p_{(1)} < p_{(2)} < \cdots < p_{(k)} < 1$

$$\lim_{n \to \infty} P(\{\prod_0^k n_i > 0\} \cap \{p(x) \text{ has } k \text{ roots} \geq 0\}) = 1. \tag{3.6}$$

Thus, for large samples, we expect to be in the case considered above, and we expect to be able to identify the MLE by equations (3.4). We briefly sketch a proof of (3.6). The fact that

$$P(\prod_0^k n_i > 0) \to 1$$

is clear from the Bonferroni inequality, since

$$P(\bigcap_{i=0}^k \{n_i \neq 0\}) \geq 1 - \sum_{i=0}^k P(n_i = 0)$$

$$= 1 - \sum_{i=0}^k (1-P(Y=i))^n .$$

This latter expression clearly tends to one as n tends to infinity. Since

$$P(p(x) \text{ has } k \text{ roots} \geq 0 | \prod_0^k n_i > 0) - P(p(x) \text{ has } k \text{ roots} \geq 0)$$

tends to zero as n tends to infinity, it suffices to show that

$$P(p(x) \text{ has } k \text{ roots} \geq 0) \to 1. \tag{3.7}$$

To see this, we focus on the polynomial

$$f(x) = \begin{cases} \frac{1}{n_0} p(x) & \text{if } n_0 > 0 \\ 1 & \text{if } n_0 = 0. \end{cases}$$

The coefficients of $f(x)$ are consistent estimates of the elementary symmetric functions of the ratios $\{\theta_{(i)} = \frac{p_{(i)}}{1-p_{(i)}},\ i=1,\ldots,k\}$. Since $P(n_0 > 0) \to 1$, we have that for each fixed x, $f(x)$ will converge in probability to the polynomial with roots $\theta_{(1)},\ldots,\theta_{(k)}$. If $p_{(1)} < p_{(2)} < \cdots < p_{(k)}$, these k roots are distinct. It is thus possible to choose points x_i, $i=0,\ldots,k$ such that $x_{i-1} < \theta_{(i)} < x_i$, and the limiting polynomial takes alternating signs at successive x's. We may choose N sufficiently large so that $f(x)$ has alternating signs at successive x's with arbitrarily high probability, establishing (3.7) which implies (3.6).

It is not possible to obtain the same result if the parameter vector is on the boundary of the parameter space. In that case, the limiting polynomial mentioned above has some roots of multiplicity greater than one. It is possible that a sequence of polynomials converges to such a polynomial, and yet no polynomial in the sequence has any real roots. For example, the polynomials $\{g_n(x) = x^4 - (6 - \frac{1}{n})x^3 + 13x^2 - 12x + 4\}$ have no real roots, yet converge to the polynomial $g(x) = (x-1)^2(x-2)^2$. It is a fortunate fact, however, that a sequence of random polynomials does not behave like a sequence of deterministic polynomials. Thus for large n, we find that the MLE may be identified from the roots of the polynomial (3.5) with reasonable frequency even in the case of a boundary parameter vector. The simulation results summarized in the next section will make this remark clearer. Thus, although the method proposed here does not succeed in identifying the MLE with limiting probability one for boundary parameter vectors as it does for vectors in the interior of the parameter space, the method may still be attempted and will produce the MLE with some positive probability -- the exact value of which depends on the exact form of the limiting polynomial.

While it is true that $P(\prod_0^k n_i > 0)$ tends to 1 as $n \to \infty$, provided $0 < p_{(1)} \leq p_{(k)} < 1$, the speed of this convergence will depend on the exact size of the parameters. There are cases of practical importance in which maximum likelihood estimation is of interest for sample sizes for which $P(\prod_0^k n_i = 0)$ is quite high. Of particular importance are problems in which several p_i's are very large and/or several p_i's are very small. As Buehler (1957) has noted, such problems occur with considerable frequency in reliability experiments. It is interesting to note that the method proposed here for the case $\prod_0^k n_i > 0$ tends to work nicely for problems in which the integers with nonzero observed frequencies are consecutive -- precisely the expected data configuration for the problem of interest. We summarize below the details of the extension of the method to this problem.

Let us suppose that the observed frequencies from a sample $Y_1,\ldots,Y_n$ are as follows:

$$n_0 = \cdots = n_{r-1} = 0 = n_{s+1} = \cdots = n_k$$

$$\text{with} \qquad \prod_{i=r}^{s} n_i > 0, \tag{3.8}$$

where $0 \leq r < s \leq k$, and $n_{-1} \equiv 0 \equiv n_{k+1}$. Then, an attempt to use the invariance property of MLE's to identify the MLE for $(p_{(1)},\ldots,p_{(k)})$, that is, an attempt to solve the system (3.2), yields

$$\hat{p}_{(1)} = \cdots = \hat{p}_{(k-s)} = 0$$

and

$$\hat{p}_{(k-r+1)} = \cdots = \hat{p}_{(k)} = 1,$$

with the remaining estimates being identified as

$$\hat{p}_{(j)} = \frac{\hat{\theta}_{(j)}}{1+\hat{\theta}_{(j)}} \qquad j=k-s+1,\ldots,k-r$$

where $\theta_{(k-s+1)},\ldots,\theta_{(k-r)}$ are the ordered roots of the polynomial

$$p(x) = \sum_{i=r}^{s} n_i(-1)^{i-r}x^{s-i}, \tag{3.9}$$

provided this polynomial has (s-r) nonnegative roots. For the case $k = 2$ considered in the last section, the polynomial (3.9) is linear in the two cases with a zero observed frequency and consecutive integers with non-zero frequencies (labeled Cases 4 and 6 there).

IV. DISCUSSION

Maximum likelihood estimation for the parameters of the model $\overset{k}{\underset{1}{*}} B(1,p_i)$ is a complex problem in which many different likelihood surfaces must be examined, and for which no closed form solution is possible in general. While numerical methods are always available for searching for MLE's, they tend to be quite unwieldy in multiparameter problems. In Section III, we have demonstrated that the MLE may be found with high probability from the roots of a k^{th} degree polynomial when n is large and the k parameters are distinct. This leaves, of course, the numerical problem of obtaining roots of this polynomial, but this problem is easily accomplished using standard techniques. It is a substantially simpler numerical problem than the problem of "hill-climbing" with a k-variate criterion function.

The results discussed in Section III are asymptotic in character, and it is of interest to examine the question of sample size requirements. We present below the result of a very modest simulation study -- we hope to

report on the results of a more ambitious simulation in a future note. For the present, we examine only the case k = 2 and the sample size n = 50. We conclude from our simulation that n = 50 is a "large sample" in terms of the level of probability experienced in identifying the MLE by the method of Section III.

In Table II below, we give, for different parameter values, the frequency of occurrence of Cases 1 - 8 (see Table I) in 100 samples of size 50 drawn from the convolution $B(1,p_{(1)}) * B(1,p_{(2)})$. The last column tabulates the frequency of occurrence of samples for which integers i with nonzero n_i are consecutive and the MLE could be identified by the invariance principle.

TABLE II

$p_{(1)}, p_{(2)}$	Case 1	2	3	4	5	6	7	8	MLE by Invariance
.1, .1						55	38	7	62
.1, .2						37	40	23	60
.1, .3						22	24	54	76
.1, .4						15	13	72	87
.1, .5						6	9	85	91
.1, .6						3	3	94	97
.1, .7						1	1	98	99
.1, .8						2		98	100
.1, .9				1		1		98	100
.2, .2						20	40	40	60
.2, .3						6	42	52	58
.2, .4						2	29	69	71
.2, .5							18	82	82
.2, .6							9	91	91
.2, .7							4	96	96
.2, .8								100	100
.2, .9				1				99	100
.3, .3							49	51	51
.3, .4							40	60	60
.3, .5							34	66	66

(continued on page 16)

TABLE II (Continued)

$p_{(1)}, p_{(2)}$	Case 1	2	3	4	5	6	7	8	MLE by Invariance
.3, .6							27	73	73
.3, .7							5	95	95
.3, .8								100	100
.3, .9				2				98	100
.4, .4							48	52	52
.4, .5							44	56	56
.4, .6							36	64	64
.4, .7							22	78	78
.4, .8							11	89	89
.4, .9				3				97	100
.5, .5							44	56	56
.5, .6							43	57	57
.5, .7							33	67	67
.5, .8							24	76	76
.5, .9				6			8	86	92
.6, .6							40	60	60
.6, .7							44	56	56
.6, .8				2			27	71	73
.6, .9				13			16	71	84
.7, .7				2			56	42	44
.7, .8				4			42	54	58
.7, .9				27			20	53	80
.8, .8				12			40	48	60
.8, .9				37			37	26	63
.9, .9				67			30	3	70

With a sample of size 50 in a two-component system, the likelihood of obtaining the MLE by the invariance principle ranges (in our simulation) from 44% to 100%, the higher likelihoods being associated with parameter values that are fairly well separated. We see that the invariance principle is not highly reliable when $p_{(1)} = p_{(2)}$, as might be anticipated by our remarks in the previous section. However, there is a reasonable chance of obtaining the MLE by invariance even in this case. For fixed $p_{(1)} = p_{(2)}$,

this likelihood should be about the same regardless of the sample size. One further observation -- while our simulation did not involve $p_{(1)} < .1$ or $p_2 > .9$, it is clear that the relative frequency with which the MLE may be obtained by invariance tends to one as either $p_{(1)} \to 0$ or $p_{(2)} \to 1$, since in either of these circumstances, the probability of the set of cases (1, 2, 3, 4, 6, 8) in which the invariance principle works tends to one.

REFERENCES

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference Under Order Restrictions. Wiley, New York.

Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing - Probability Models. Holt, Rinehart and Winston, New York.

Blischke, W. R. (1964). Estimating the parameters of mixtures of binomial distributions. Journal of the American Statistical Association, **59**, 510-28.

Buehler, R. J. (1957). Confidence intervals for the product of two binomial probabilities. Journal of the American Statistical Association, **52**, 482-93.

Gaffey, W. R. (1959). A consistent estimator of a component of a convolution. Annals of Mathematical Statistics **30**, 198-205.

Samaniego, F. J. (1976). A characterization of convoluted Poisson distributions with applications to estimation. Journal of the American Statistical Association, **71**, 475-79.

Samaniego, F. J. (1977). Maximum likelihood estimation for binomially distributed signals in discrete noise. Journal of the American Statistical Association, to appear.

Sclove, S. L. and Van Ryzin, J. (1969). Estimating the parameters of a convolution. Journal of the Royal Statistical Society, B **31**, 181-91.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) Maximum Likelihood Estimation for a Class of Multinomial Distributions Arising in Reliability.		5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) F. J. Samaniego ~~and~~ L. E. Jones		8. CONTRACT OR GRANT NUMBER(s) (15) AFOSR-77-3180
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of California Department of Mathematics Davis, California 95616		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16) (17) 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, D.C. 20332		12. REPORT DATE (11) July 1978
		13. NUMBER OF PAGES 18
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 21p. (14) TR-6		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

(18) AFOSR

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(19) TR-78-1267

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Convolution, maximum likelihood estimation, invariance principle, reliability.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let X_i, $i=1,\ldots,k$ be independent Bernoulli variables, with $X_i \sim B(1,p_i)$. Let $Y = \Sigma X_i$, and consider a multinomial experiment based on n independent, identically distributed observations $Y_1,\ldots,Y_n$. This model is identifiable in the ordered parameter vector $(p_{(1)},\ldots,p_{(k)})$, where $p_{(i-1)} \le p_{(i)}$, and arises in reliability experiments in which k components in parallel have potentially

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

401 891

LB

20. CONTINUED...

different probabilities of failure. The family of multinomial distributions with the structure described above is properly contained in the family of general multinomial distributions with (k+1) classes. Maximum likelihood estimation for this family is considered, and it is shown that for sufficiently large n, the maximum likelihood estimate of $(p_{(1)},\ldots,p_{(k)})$ may be identified with high probability from the roots of a k^{th} degree polynomial whose coefficients are consistent estimates of the elementary symmetric functions of the ratios $\theta_{(i)} = p_{(i)}/(1-p_{(i)})$. A simulation study for the case $k = 2$ sheds light on the sample size required.